



# Individuality and slow dynamics in bacterial growth homeostasis

Lee Susman<sup>a,b</sup>, Maryam Kohram<sup>c</sup>, Harsh Vashista<sup>c</sup>, Jeffrey T. Nechleba<sup>c</sup>, Hanna Salman<sup>c,d,1</sup>, and Naama Brenner<sup>a,e,1</sup>

<sup>a</sup>Network Biology Research Laboratory, Technion—Israel Institute of Technology, 32000 Haifa, Israel; <sup>b</sup>Interdisciplinary Program of Applied Math, Technion—Israel Institute of Technology, 32000 Haifa, Israel; <sup>c</sup>Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260; <sup>d</sup>Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15260; and <sup>e</sup>Department of Chemical Engineering, Technion—Israel Institute of Technology, 32000 Haifa, Israel

Edited by Stanislas Leibler, The Rockefeller University, New York, NY, and approved May 7, 2018 (received for review September 19, 2016)

**Microbial growth and division are fundamental processes relevant to many areas of life science. Of particular interest are homeostasis mechanisms, which buffer growth and division from accumulating fluctuations over multiple cycles. These mechanisms operate within single cells, possibly extending over several division cycles. However, all experimental studies to date have relied on measurements pooled from many distinct cells. Here, we disentangle long-term measured traces of individual cells from one another, revealing subtle differences between temporal and pooled statistics. By analyzing correlations along up to hundreds of generations, we find that the parameter describing effective cell size homeostasis strength varies significantly among cells. At the same time, we find an invariant cell size, which acts as an attractor to all individual traces, albeit with different effective attractive forces. Despite the common attractor, each cell maintains a distinct average size over its finite lifetime with suppressed temporal fluctuations around it, and equilibration to the global average size is surprisingly slow (>150 cell cycles). To show a possible source of variable homeostasis strength, we construct a mathematical model relying on intracellular interactions, which integrates measured properties of cell size with those of highly expressed proteins. Effective homeostasis strength is then influenced by interactions and by noise levels and generally varies among cells. A predictable and measurable consequence of variable homeostasis strength appears as distinct oscillatory patterns in cell size and protein content over many generations. We discuss implications of our results to understanding mechanisms controlling division in single cells and their characteristic timescales.**

biophysics | microorganism populations | single-cell measurements | bacterial growth | phenotypic variability

The processes of growth and division in proliferating cells have been of interest for decades, with microorganisms providing model systems for both experimental and theoretical studies. Recently, with the development of experimental methods (1), new light was shed on this problem. Large samples of cells can be continuously tracked as they grow and divide for multiple cycles at high spatial and temporal resolution. Such measurements provide ample new information about these processes. As an important quantitative result of single-cell tracking, it was shown that individual yeast cells (2) as well as different bacterial cells (3–6) grow exponentially in time to a good approximation between consecutive divisions. This result sharpens the problem of cell size homeostasis, since successive cycles of exponential growth and division can be unstable to fluctuations (7–11). Negative correlations can, in principle, prevent the instability and divergence caused by independent fluctuations. Indeed, using large samples of growth and division cycles pooled from single-cell measurements, a negative correlation was found between the size change over the cell cycle and the initial size (i.e., cell size at the start of the cell cycle) (12–17). The observed correlations, if interpreted as regulation of cell division, rule out two previously studied models of division control, namely a constant fold change over the cell cycle and a fixed size threshold for division. However, despite these new data and insights, a

clear mechanism linking cell size to division has not yet been identified (18).

From a theoretical modeling perspective, the experimentally observed correlations provide the basis for a phenomenological approach without reference to any specific underlying mechanism (7, 8, 19–21). Such an approach can be formulated mathematically in several essentially equivalent ways (20), all of which incorporate the exponential accumulation and its above-mentioned negative correlation with the initial cell size. Although this correlation does not necessarily imply a causal relation, it is convenient to envision it as a restraining force that counteracts the fluctuating exponentials, attracting them to the bottom of an effective potential well (20, 21). In this picture, a cell born too large will accumulate a smaller fold change over the growth cycle and vice versa, preventing fluctuations from accumulating in the long run.

Several experiments in *Escherichia coli* and *Saccharomyces cerevisiae* were found to be consistent with a specific value of the effective force constant. This particular value corresponds, in a linear approximation, to a fixed volume added on average at each cycle and was hence termed the “adder” model (8, 13, 14, 22). Closer inspection of the data, however, reveals that the correlation plots are very noisy, despite the large samples and high accuracy of the experiments. Moreover, some experiments showed force constants different from the one corresponding to the adder model. Investigation of *E. coli* and mycobacteria in different environments, for example, resulted in a range of different

## Significance

**Microbial cells go through repeated cycles of growth and division. These cycles are not perfect: the time and size at division can fluctuate from one cycle to the next. Still, cell size is kept tightly controlled, and fluctuations do not accumulate to large deviations. How this control is implemented in single cells is still not fully understood. We performed experiments that follow individual bacteria in microfluidic traps for hundreds of generations. This enables us to identify distinct individual dynamic properties that are maintained over many cycles of growth and division. Surprisingly, we find that each cell suppresses fluctuations with a different strength; this variability defines an “individual” behavior for each cell, which is inherited along many generations.**

Author contributions: H.S. and N.B. designed research; L.S., M.K., H.V., J.T.N., H.S., and N.B. performed research; L.S., M.K., H.V., J.T.N., H.S., and N.B. analyzed data; and L.S., H.S., and N.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

<sup>1</sup>To whom correspondence may be addressed. Email: [hsalman@pitt.edu](mailto:hsalman@pitt.edu) or [nbrenner@technion.ac.il](mailto:nbrenner@technion.ac.il).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615526115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615526115/-DCSupplemental).

Published online June 5, 2018.

measured values (15, 16). Experiments in the *Caulobacter crescentus* revealed two phases in the cell cycle, each characterized by a different restraining force strength (17).

All of these previous studies have used measurements pooled from many single cells to increase statistics. Such pooled data can provide information on cell cycle parameters averaged over the entire ensemble of cells; however, mechanisms that regulate and control division operate at the level of the single cell, and their individual properties might be masked by such pooling. In this study, we measure and analyze dynamics of growth and division in individual bacteria tracked over extended times: up to ~250 cell cycles each. Making the distinction between statistics over time in individual cells and the corresponding statistics averaged over many cells requires, first, that long enough stable individual traces be acquired without confounding effects, such as filamentation or contamination, and second, that statistical properties be analyzed from separate traces and compared with those averaged over many traces. Our previous work carried out such a comparison for protein distributions and found a universality of distribution shape in both ensembles (12). To make a similar comparison for cell size homeostasis, which is a dynamic process, longer traces and more statistics are required. Here, we present data that enable this comparison.

Our results show that individual cells exhibit different values of the effective restraining force constant, which is maintained distinct for many cell cycles. At the same time, an invariant is revealed in the form of an ensemble average cell size, acting as an attractor to the dynamics over long times. Despite this common attractor, we find significant differences in temporally averaged size between traces over the finite lifetime of each cell. This is related to deviations of temporally averaged division ratio and fold change from their global average values of 1/2 and 2, respectively. Such deviations are persistent over dozens of division cycles, and equilibration to the global averages appears only in the longest traces: those over 150 cycles long.

Integrating cell size data with measurements of protein content in the same cells, we propose that a possible origin of variable homeostasis strength stems from underlying interactions between global cellular variables. We present an illustrative mathematical model of these interactions, which reproduces several nontrivial aspects of the entire dataset. As a consequence of the individuality in the homeostasis parameter (restraining force constant), we provide a theoretical explanation for oscillatory autocorrelations in cell size and in protein content, which have been previously reported (15). We discuss the implications of our results to the quest for the mechanisms underlying cellular growth and division homeostasis and point to future research directions.

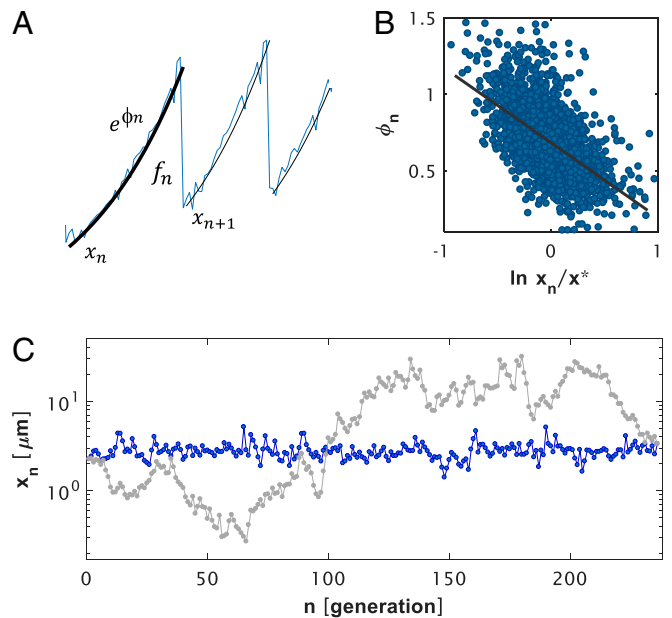
## Results

### Cell Size Homeostasis: Single-Cell vs. Ensemble Average Behavior.

Continuous measurements of cell size over time reveal smooth, exponential-like accumulation throughout each cell cycle interrupted by abrupt drops at division. Fig. 1A shows a small portion of such a measurement. Cell length is taken as an attribute of cell size, as the rod-like *E. coli* bacteria grow in one dimension along their length, while their width is maintained constant (*Methods*) (3, 12). Over the  $n$ th cycle of growth and division, cell size  $x_n(t)$  can be described accurately as

$$\begin{aligned} x_n(t) &= x_n(0)e^{\alpha_n t}, & 0 < t < T_n \\ x_{n+1}(0) &= f_n x_n(T_n), \end{aligned}$$

where  $\alpha_n$  is the exponential accumulation rate during cell cycle  $n$ ;  $T_n$  is its duration, and  $f_n$  is the division fraction at its end (black fitting line in Fig. 1A). The cell size at the start of the  $n$ th cell



**Fig. 1.** Correlations in cycles of exponential accumulation and division. (A) A portion of a trace measuring the size of a trapped bacterium along time, illustrating the exponential accumulation within cycle  $n$ ,  $e^{\phi_n}$ , and the division fraction,  $f_n$ . These two variables connect the initial cell size  $x_n$  with that at the next cycle  $x_{n+1}$  (Eq. 1); both fluctuate from one cycle to the next. (B) Exponential accumulation  $\phi_n$  is negatively correlated with  $\ln x_n$  (best fit slope for Eq. 2:  $\beta = 0.49 \pm 0.02$ ). Taking  $x^*$  to be the average cell size,  $x^* = 2.7 \mu\text{m}$ , we find  $\phi^* = 0.69 \pm 0.2$ . (C) The blue line indicates a long sequence of initial cell sizes  $x_n$  from one trace as a function of cycle number (generation). The gray line indicates a shuffled process created from the measured pairs  $(e^{\phi_n}, f_n)$  by applying them as fold changes to the initial condition of the trace in random order, thus discarding the correlation between fold change and initial cell size.

cycle  $x_n(0)$ , which we denote simply as  $x_n$ , is linked across generations by the mapping

$$x_{n+1} = f_n x_n e^{\phi_n}, \quad [1]$$

with the total accumulation exponent defined as  $\phi_n = \alpha_n T_n$  (7). In this discrete mapping, each step represents a complete cell cycle and is characterized by two variables: a total increase in cell size by a factor  $e^{\phi_n}$  from the beginning to the end of the cycle and a decrease by a factor of  $f_n$  at division. The definition of these parameters is illustrated in Fig. 1A. Both increase and decrease are subject to fluctuations from one cell cycle to the next (*SI Appendix, Fig. S1-1* shows their distributions); cell size homeostasis requires their product to average to one over long times. Pooling together a large sample of cell cycles from many individual traces shows that, on average, this is indeed the case (*SI Appendix, Fig. S1-1*). However, while this requirement is necessary, it is insufficient for homeostasis: the process described by Eq. 1 is unstable against fluctuations in  $f_n$  and  $e^{\phi_n}$  over long times, even if, on average, their product is one; independent fluctuations accumulate, and the variance increases with time. Fig. 1B shows that the exponential accumulation of size during a cell cycle,  $\phi_n$ , is negatively correlated with initial cell size. Such state-dependent changes can control fluctuations and induce a stable size distribution over multiple generations (7, 8). Additional empirical correlations are presented in *SI Appendix, Fig. S1-2*.

To show the effect of state-dependent changes on long-term dynamics, we present in Fig. 1C a comparison between two sequences of cell sizes  $x_n$ : the first (Fig. 1C, blue) is an actual measurement of a single cell for over 200 generations, while the second

(Fig. 1C, gray) is created artificially by randomly shuffling the order of the pairs  $(\phi_n, f_n)$  in the same trace. Such a permutation retains the average of the products  $e^{\phi_n} f_n$  at one but loses the correlations displayed in Fig. 1B. Clearly, these correlations contribute to the stabilization of cell size, limiting it to a small range.

One may account for the negative correlations in Fig. 1B in the mapping (Eq. 1) by postulating a relation between  $x_n$  and the fold change  $e^{\phi_n}$  (7, 8). Consistent with the data, we may use a linear approximation in logarithmic coordinates:

$$\phi_n = \phi^* - \beta \ln \frac{x_n}{x^*} + \xi_n, \quad [2]$$

where  $\beta$  represents the slope of the fit in Fig. 1B. We term this the homeostasis parameter. The noise,  $\xi_n$ , has zero mean and is approximately Gaussian. Here, the typical size  $x^*$  sets the units in which cell size is measured. This scaling size  $x^*$  is chosen to be the average cell size at the start of the cell cycle over the entire dataset or the “ensemble average” cell size. We find that the best linear fit is obtained with  $\phi^* \approx \ln 2$ , corresponding to a mean fold change of 2, as expected, when averaging over many cells with a mean division ratio of 1/2 (SI Appendix, Fig. S1-1).

We next consider the same analysis of correlations between cell cycle variables applied to individual traces separately; a similar picture may be expected but with some degree of variability among individual cells. This variability could be due to noise, in which case they will have similar linear correlation parameters up to errors resulting from measurement noise and finite sampling. However, significant variability in the correlations could reflect true individuality of cells. We write the analog of Eq. 2 for the  $k$ th individual trace,

$$\phi_n^{(k)} = \phi^{*(k)} - \beta^{(k)} \ln \frac{x_n^{(k)}}{x^*} + \xi_n^{(k)}, \quad [3]$$

with cell size still measured in units of the ensemble average  $x^*$ . Variability can be reflected as significant differences in any one of the parameters  $\phi^{*(k)}$  or  $\beta^{(k)}$  or the properties of the noise  $\xi_n^{(k)}$ .

In graphical terms, the pooled scatterplot of Fig. 1B could be composed of single-cell plots that differ in their properties in several alternative ways; however, they could also exhibit invariant features common to all traces (illustrations are in Fig. 2). Identifying those properties that are conserved among all cells may point to their importance as control variables.

Fig. 3A highlights two individual measured traces in color (Fig. 3A, green and red) on the background of the entire ensemble in gray, suggesting that they have distinct values of the homeostasis

parameter. Fig. 3B shows the estimated slopes  $\beta^{(k)}$  for all traces as a function of trace length in number of generations. The error bars, representing the uncertainty in the slope, are significantly smaller than the differences between slopes for traces of length <150 generations; these differences decrease for the extremely long traces of length 150–250 generations. To quantify the difference between traces, two methods of statistical analysis were applied, showing that the apparent difference is statistically significant beyond the noise and the finite sampling (SI Appendix).

Fig. 3C shows the best linear fits for all individual traces. Each black line was obtained as a fit similar to the colored lines in Fig. 3A taken from a single individual trace along time. This figure reveals an invariant in the form of a pivot point where all lines cross. The coordinates of this point coincide with the ensemble average (green circle in Fig. 3C), implying a common intercept  $\phi^{*(k)} = \phi^* = \ln 2$  for all traces. This leaves one parameter, the homeostatic parameter or correlation slope  $\beta^{(k)}$ , which is distinct to each trace.

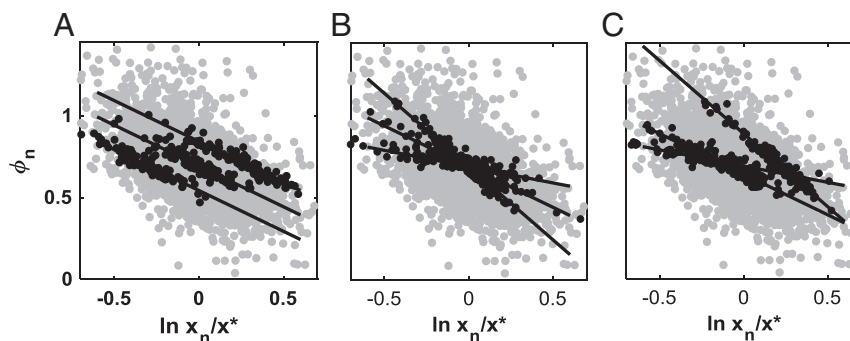
The pivot point, common to all individual traces, suggests a dynamic attractor for cell size over multiple cycles. In this picture, if the cell divides to a size that considerably deviates from  $x^*$ , the exponential accumulation during the following cell cycle would be compensated to effectively “pull” the cell back to this common attractor, with a force strength variable among cells. This picture is supported by Fig. 3D, in which a flow map computed as an average over all cycles in the dataset is presented. Similar dynamics are also found in other experimental conditions (for example, a different nutrient composition) (SI Appendix, Fig. S4). We note that, while  $\phi^* = \ln 2$  for all conditions tested, the ensemble average cell size at the start of the cell cycle,  $x^*$ , depends on growth medium and temperature.

#### Differences Between Time-Averaged Cell Sizes of Individual Traces.

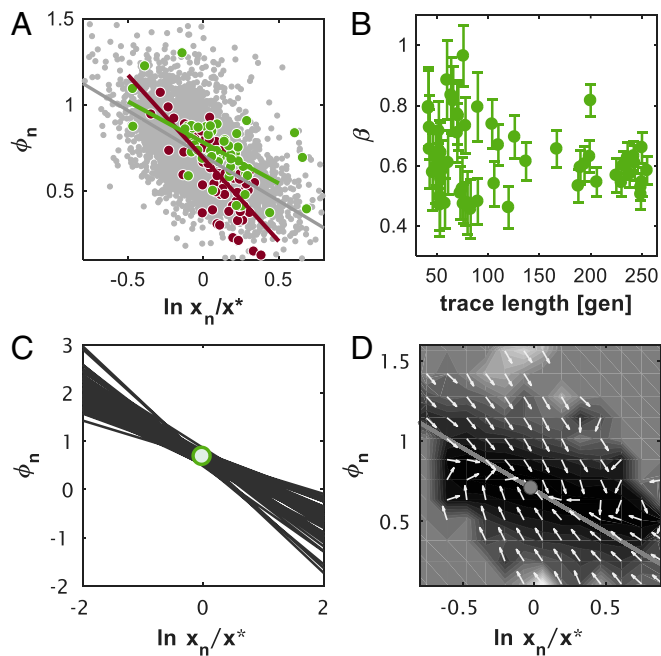
Do the distinct values of  $\beta^{(k)}$  result in different cell sizes when averaged over the lifetime of the cell? We use the mapping model for individual traces to answer this question: combining Eqs. 1 and 3, the mapping can be written as

$$\ln \frac{x_{n+1}^{(k)}}{x^*} = (1 - \beta^{(k)}) \ln \frac{x_n^{(k)}}{x^*} + \ln f_n^{(k)} + \phi^* + \xi_n^{(k)}, \quad [4]$$

linking the logarithm of cell size in consecutive cell cycle starts. Here, we have incorporated the empirical observations that homeostasis parameters are distinct, whereas the intercept  $\phi^*$  is common to all traces. Denoting temporal averaging over a trace



**Fig. 2.** Possible patterns of variability and invariants in single-cell trace correlations. (A) Single-cell traces could exhibit the same slope of correlation, indicating that the effective restraining force strength is a relevant control variable. (B) Alternatively, they could exhibit an invariant intersection point, pointing to a preferred common cell size. (C) Traces could also be variable in both properties without conserving any global invariant. In all panels, gray dots are measurement data. Black dots result from simulating the mapping model (Eqs. 1 and 2) with different slopes and intercepts  $(\beta, \phi^*)$  for the correlation of Eq. 2. In all cases, the average division ratio is 1/2, and the average fold change over the trace is 2.



**Fig. 3.** Individual (A and B) and common (C and D) aspects of cell size homeostasis in bacterial traces. (A) The same data as in Fig. 1B are plotted in gray for the entire ensemble of traces. Points from two individual traces are highlighted in color with their respective best linear fits, displaying a different slope for each and thus, a different homeostasis parameter  $\beta$ . (B) Estimated slopes  $\beta$  for all individual traces as a function of their length in number of generations [gen], with error bars denoting the SE in the estimate (Methods). (C) Best linear fits for all individual traces intersect at a common pivot point. The green circle indicates ensemble average of the two axes. (D) A flow map is estimated in the 2D phase space  $(\ln x_n/x^*, \phi_n)$ . The flow direction is indicated by arrows; its amplitude is encoded in the underlying heat map [contours have a uniform spacing of 0.1 and range from 0 (black) to 2 (white)]. The pivot point of C (gray circle) is an attractor on this projection of the dynamics.

by overbars, we compute from Eq. 4 the time-averaged logarithm of initial cell size

$$\overline{\ln x_n^{(k)}} = \ln x^* + \frac{\phi^* + \overline{\ln f_n^{(k)}}}{\beta^{(k)}}, \quad [5]$$

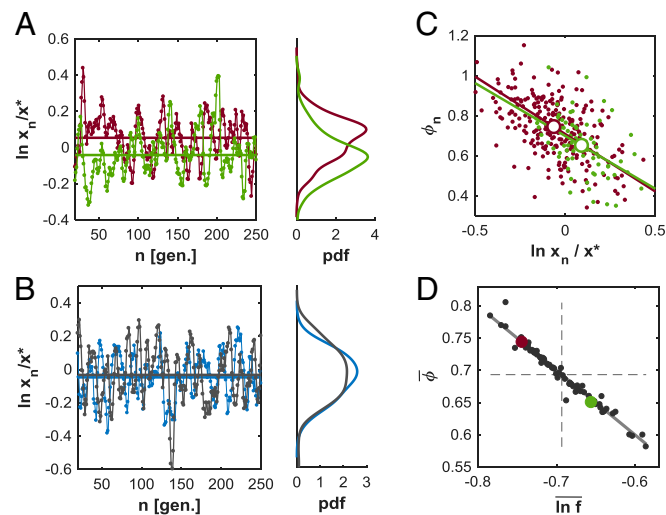
where we have used the empirical result  $\overline{\xi_n^{(k)}} \approx 0$ . If  $\overline{\ln f_n^{(k)}} = \ln 1/2$ , the right term vanishes:  $\phi^* + \ln f_n^{(k)} = \ln 2 + \ln 1/2 = 0$ ; then, distinct values of  $\beta^{(k)}$  may affect the rate of relaxation toward the attractor but not the steady state itself. However, when we examine experimentally measured sequences of consecutive initial cell sizes  $x_n$ , we find that their averages are distinct. Fig. 4A displays such sequences for two long and stable traces, with horizontal lines depicting their temporal averages. The distribution of values along the trace is plotted in Fig. 4A, Right for each trace with its corresponding color.

To understand the origin of these differences, we simulated two traces using the model, Eq. 4, with homeostasis parameters  $\beta^{(k)}$  ( $k=1, 2$ ) taken from the two traces in Fig. 4A and with  $\phi^* = \ln 2$ . Both  $\xi_n$  and  $f_n$  were simulated as random variables drawn independently at each step, with statistical properties matching those of the ensemble (i.e.,  $\xi_n = 0$  and  $\langle \ln f_n \rangle = \ln 1/2$ ). Fig. 4B shows the two simulated traces. In comparison, the measured traces in Fig. 4A exhibit suppressed temporal fluctuations, each around a different mean value, farther removed from one another than the model predicts. These effects can be quantified by computing the SD “internal” to an

individual trace,  $\sigma_{in}(\ln(x/x^*))$ , estimated over time (width of distributions in Fig. 4A, Right and B, Right). We find that, on average over all measured traces,  $\langle \sigma_{in}(\ln(x/x^*)) \rangle = 0.26 \pm 0.06$ , while for a corresponding collection of simulated traces, we find  $0.43 \pm 0.05$ . However, fluctuations of the measured traces are centered around temporal averages, which are significantly different from one another. This can be quantified by the “external” SD of time-averaged sizes,  $\overline{\ln(x/x^*)}$ , across all traces: we find that  $\langle \sigma_{ex}(\overline{\ln(x/x^*)}) \rangle = 0.14$  for experimental data and 0.06 for the model simulation results (details are in Methods). This analysis provides statistical support to the effect apparent in Fig. 4A, namely that cell size fluctuations along time are strongly suppressed in each trace around a distinct time-averaged value.

The discrepancy between individual traces and the model prediction suggests a distinct behavior of the division ratio in each trace: in the model, this was taken as a random variable common to all traces and drawn from a Gaussian distribution around 1/2. However, each cell undergoes a limited number of growth and division cycles before it dies. Therefore, if division ratios maintain a bias that deviates on average from 1/2 along many cycles, the effective feedback in the exponential accumulation  $\phi_n$ —which ensures balanced growth—induces a corresponding deviation of the average fold increase from 2. Consequently, the range of values sampled in the  $(\ln x_n, \phi_n)$  plane by an individual trace over its finite lifetime may be biased and may not provide a good sample of the range attained by the entire ensemble.

This signature of slow dynamics manifests as a distinct clustering of the points making up each trace as illustrated in Fig. 4C. The clusters corresponding to the two individual traces, indeed,



**Fig. 4.** Individuality and slow dynamics in cell size traces. (A, Left) Two measured individual traces (colors) showing cell size in consecutive cell cycle starts,  $x_n$  (three-point smoothed) as a function of generation number [gen]. Horizontal lines indicate time averages of each trace. (A, Right) Probability density functions (pdfs) of cell size values for the two plotted traces (corresponding colors). Each cell maintains fluctuations around a distinct mean value, with internal SD  $\langle \sigma_{in}(\ln x) \rangle = 0.26$  averaged over all traces. The SD of temporal averages among traces is 0.14. (B) The same as in A for two simulated traces (Eq. 4), with parameters matching those in A. The internal SD of simulated traces,  $\langle \sigma_{in}(\ln x) \rangle = 0.43$ , is larger for model traces. The SD of the temporal average across traces is  $\langle \sigma_{ex}(\overline{\ln x}) \rangle = 0.06$ . (C) Two measured traces exhibit distinct clustering in the  $(\ln x_n, \phi_n)$  plane (colors). As a consequence, each cell maintains a distinct average size (large circles) over its lifetime. (D) Temporal averages of division ratio and accumulation exponents in all measured traces. The solid line indicates  $\overline{\phi_n} + \ln \overline{f_n} = 0$ . The two traces shown in C are highlighted in color.

are not overlapping, and each trace samples a slightly different portion of the space. Large colored circles in Fig. 4C depict the average of the colored points corresponding to two individual traces and illustrate how the distinct clusters result in distinct average sizes. The time-averaged division ratios and the corresponding time-averaged exponential accumulations are presented for all traces in Fig. 4D, showing the extent of biases spanned by individual traces and the tight compensation between them induced by homeostasis. This explanation of distinct time-averaged sizes in terms of division ratio bias agrees reasonably well with the data as detailed in *SI Appendix, Fig. S5*.

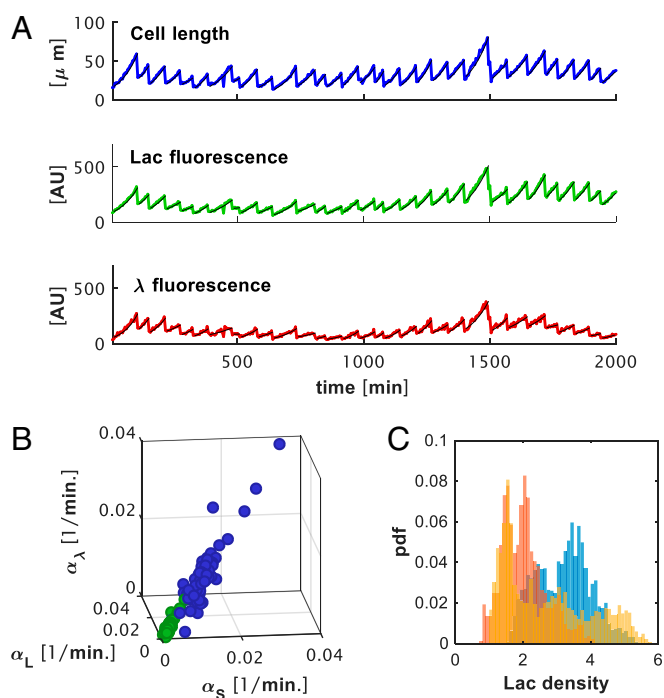
**Possible Origin of Variation in Homeostasis Parameters.** Why do individual cells exhibit distinct values of the homeostasis parameter,  $\beta^{(k)}$ ? Recall that this parameter quantifies an empirical negative correlation between initial cell size and exponential size accumulation during the cell cycle. Mechanistically, little is known about underlying processes that may induce such a negative correlation. One might imagine that some molecular circuit implements a feedback loop from accumulated cell size to division; experiments have suggested different molecules to be implicated in such a process, but a specific mechanism has not yet been identified (18). Recently, several researchers have put forward the possibility that size homeostasis is not implemented at the molecular level but may represent a global systems-level property of the cell (18, 23–25).

In line with this idea, we consider the dynamics of accumulation and division from a global cellular perspective. A key observation is that the copy number of highly expressed proteins, which also accumulates exponentially and divides over multiple cell cycles, exhibits an apparent “protein homeostasis,” reflected in negative correlations similar to cell size; namely, one may associate a nonzero value of  $\beta$  with highly expressed proteins in the cell (7). These values, which are generally smaller than those corresponding to cell size, span a wide range for different proteins, conditions, and individual cells (*SI Appendix, Fig. S6*).

Previous work has shown that traces of highly expressed protein content are not only qualitatively similar to those of cell size but also, statistically correlated with them on a cycle by cycle basis (12). To further characterize this relationship, we measured the copy numbers of two fluorescent proteins simultaneously in single *E. coli* cells along cycles of growth and division. One was expressed under the control of the *lac* promoter in a lactose-rich medium and thus, represents a metabolically relevant protein. The second was expressed under the control of the constitutive  $\lambda$ -phage  $p_R$  promoter ( $\lambda$ - $p_R$ ), which is foreign to the cell and does not contribute to cellular metabolism (details are in *Methods*). Examples of three simultaneously measured traces can be seen in Fig. 5A. A strong correlation between the three exponential rates measured in the same cell cycle is seen in Fig. 5B.

Given these similarities and quantitative correlations, one may argue that the copy number of protein is simply proportional to cell size. This would imply a constant (or narrowly distributed) protein density per unit volume. However, Fig. 5C shows that protein density spans a broad range, approximately fivefold in concentration, suggesting that the relationship between protein and cell size is not a simple proportion. Furthermore, a model of several phenotype components that are “enslaved” to cell size, accumulating with the same exponential rates and dividing at the same times controlled by cell size, is found to be unstable to fluctuations and cannot induce homeostasis on the entire multidimensional system (*SI Appendix, Fig. S7-2*).

If protein content is not enslaved to cell size and is not directly implicated in cell division control, why then does it have nonzero effective restraining force strength  $\beta$ ? Taking a holistic view on cellular homeostasis, we consider the possibility that effective interactions between various measurable cellular characteristics



**Fig. 5.** Relationship between cell size and content of highly expressed proteins. (A) Traces of cell size (Top) and two highly expressed proteins [fluorescent proteins expressed from the *lac* promoter (Middle) and the  $\lambda$ - $p_R$  promoter (Bottom)]. All three components exhibit cycles of exponential accumulation and division. (B) Exponential accumulation rates of cell size ( $\alpha_S$ ), *lac* expression ( $\alpha_L$ ), and  $\lambda$ - $p_R$  expression ( $\alpha_\lambda$ ) are strongly correlated across cycles. Each dot in the 3D space represents the three exponential rates corresponding to one cell cycle. (C) Protein density in three individual cells (different colors), each collected over multiple cycles of growth and division, displays a broad distribution. pdf, probability density function.

(cell size, protein content, etc.) give rise to an effective homeostasis parameter for protein content. We shall see that such a model also explains the variation among cells in the homeostasis parameter corresponding to cell size.

As a concrete implementation of this principle, we consider a set of  $D$  cellular components, which we denote by a vector  $\vec{x}$ . To describe the effective interactions, we go beyond the mapping model, which only relates discrete time points in consecutive generations, and also include dynamics of components within the cell cycle. Building on previous models of linear interactions, which give rise to indirectly autocatalytic dynamics of all variables (26, 27), we write the equation of motion within cell cycle  $n$  as

$$\frac{d}{dt} \vec{x}_n(t) = \mathcal{K} \vec{x}_n(t), \quad 0 < t < T_n, \quad [6]$$

where  $\vec{x}_n(t)$  now describes the continuous evolution of all  $D$  components. The effective interaction matrix  $\mathcal{K}$  is randomly chosen and describes intracellular metabolism fixed along cycles. Cell division distributes fractions  $f_{j,n}$  and  $(1 - f_{j,n})$  of component  $j$  to each daughter cell,

$$x_{j,n+1}(0) = f_{j,n} \cdot x_{j,n}(T_n), \quad [7]$$

with  $f_{j,n}$  randomly distributed around 1/2. The model description is completed by designating component 1 as controlling cell division through the relation in Eq. 2 (other control strategies are in *SI Appendix*). Fig. 6A shows traces of three components in the same cell resulting from numerically simulating the model. It



model, we now show that the structure of these correlations is the predictable outcome of variable values of  $\tilde{\beta}_j^{(k)}$  in individual cells and in different phenotype components (cell size, protein).

For any given  $\beta$ , computing the ACF via Eq. 4 by averaging over the ensemble from which noise is drawn, one finds an exponential with a time constant of  $\approx \ln(1-\beta)$  (in number of generations). This is in line with the smooth form that appears after averaging over all traces, where  $\beta$  is some typical value in the ensemble (Fig. 7A and B, black lines). In a single trace, however, one cannot directly calculate the ACF from the model. However, the probability of oscillatory patterns being generated at random across time and their period may be estimated (28) (SI Appendix). This probability depends on the trace-specific value of  $\beta$  and therefore, will be reflected in distinct oscillatory patterns.

Fig. 7C shows this theoretical prediction (black solid line in Fig. 7C) together with the corresponding quantities computed from our experimental traces (dots in Fig. 7C). Although the individual traces show a large scatter, binning them by value of  $\tilde{\beta}_j^{(k)}$  agrees well with the theory (large circles in Fig. 7C). The scatter is expected, since the theory is probabilistic and predicts an average over realizations; it becomes a better predictor of model simulations as trace length increases (SI Appendix, Fig. S8). We see that the oscillatory patterns of the ACF arise from purely stochastic effects in combination with the inherent discreteness of cell division and that the individuality of the homeostasis parameter echoes in their distinct periods. This agreement of the theoretical prediction with the data provides an independent verification of the variability in  $\tilde{\beta}_j^{(k)}$ , specifically in the cell size homeostasis parameter, among individual traces.

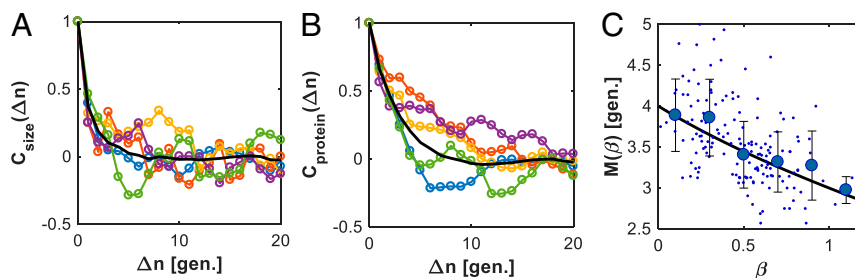
## Discussion

The process of cellular growth and division is subject to many sources of noise, which can accumulate and lead to divergence over time if left unrestrained. In an effort to understand the restraining forces that maintain cell size homeostasis in bacteria, we have analyzed the size dynamics of many individual cells measured for up to hundreds of generations. Such dynamics can be described by a phenomenological model, with an effective feedback linking the exponential size accumulation during each cell cycle,  $\phi_n$ , to the initial size in that cycle,  $x_n$ . This feedback, which is a negative correlation inferred directly from the data, acts as a restraining force for maintaining cell size from diverging over time. It has often been interpreted as a mode of division regulation, where specific restraining force strengths correspond to previously described regulation modes (e.g., adder, sizer, etc.). Different modes display different slopes of the correlation in the  $(\ln x_n, \phi_n)$  phase space.

**Cell Size Homeostasis: Pooled Cycles Vs. Individual Traces.** When analyzing the data after pooling cycles from many individual traces, the correlation is consistent with a slope of 1/2, corresponding to the previously studied adder model. However, examining the data for each cell separately reveals that the correlation slopes vary from cell to cell. This variation is statistically significant beyond the noise in the measurements. Examining the entire collection of traces in our dataset, we find that the best linear fits cross at a common point, corresponding to the average cell size and the average exponential accumulation of  $\ln 2$ . The fact that all individual lines cross at one point is a nontrivial result; in principle, they could have varied in other ways that would remain consistent with the observed ensemble scatterplot (Fig. 2). This result points to a physiologically invariant cell size acting as a common attractor of the dynamics. Thus, in individual cells, homeostasis pulls against fluctuations toward a common cell size, albeit with different force strength. The actual value of the cell size at this attractor depends on experimental conditions.

Despite the common attractor, our measurements reveal that time-averaged cell sizes remain distinct among traces over dozens of generations. The difference between temporal averages of individual traces reflects slow dynamics that extend over this timescale. In particular, the exponential accumulation and division ratio do not always converge to the ensemble averages of 2 and 1/2, respectively, over the lifetime of the cell. This may seem to be a surprising result; however, in principle, the existence of an effective feedback allows each trace to remain centered around a distinct steady-state value without losing homeostasis. The mapping model can empirically predict these deviations reasonably well as stemming from slow dynamics of division ratios with temporal averages that can deviate from 1/2 over many generations. Additional work is required to understand how long-term deviations in a mother cell are reconciled with the behavior of its daughter cells and how eventually the lineages make up a population with symmetric division on average (29). These are topics for future research.

**Homeostasis of Multiple Cellular Components.** To better understand homeostasis in individual cells, we examined the dynamics of not only their size but also, highly expressed proteins across many cycles of growth and division. Most proteins in bacteria are highly expressed, with a relatively small effect of number fluctuations (30) and with degradation negligible over the timescale of a cell cycle (31, 32). These properties result in protein content being a global cellular variable, buffered from many microscopic noise sources. Its global nature results in a universal distribution shape, insensitive to many control parameters (33). Previous work has shown that the long-term dynamics of protein accumulates exponentially within the cell cycle and exhibits effective homeostasis similar to cell size; thus, it can be described by the



**Fig. 7.** ACFs of individual traces. (A) ACFs as a function of generation lag [gen.], several cell size traces (colors). Black indicates the average over all traces. (B) ACFs for traces of fluorescent protein expressed from the  $\lambda$ -p<sub>R</sub> promoter. Black indicates the average over all traces. (C) Mean peak-to-peak distance  $M(\beta)$  in number of generations [gen.] computed from all individual traces of cell size and  $\lambda$ -p<sub>R</sub> expression. Binned data are shown as large circles. The black curve indicates the prediction from theory (28) (SI Appendix).

same mapping model applied to cell size (7, 12). This analogy suggests a strong coupling between cell size and protein.

To further characterize this coupling, we have measured cell size simultaneously with two highly expressed proteins (metabolically relevant and irrelevant) in the same cell and analyzed how these components of the phenotype coevolve over multiple generations. All components accumulate exponentially, exhibiting a strong positive correlation between the accumulation rates on a cycle by cycle basis. Nevertheless, the measured relation between them is inconsistent with a simple proportion or a dominance of one component (e.g., cell size) that determines all others up to some noise. These results suggest treating proteins and cell size as coupled components of a multidimensional interacting system. Therefore, we studied such models with different coupling schemes between the components and compared them with the integrated set of experimental results.

At the level of mapping between generations with effective feedback, it is difficult to achieve homeostasis of multiple components that exponentially accumulate and divide when one variable controls division. Perhaps surprisingly, this is true even if their exponential accumulation rates are identical up to a reasonable noise level. We found that a simple way to induce such homeostasis is by including the dynamic coupling between components during the cell cycle. Random linear interactions were sufficient to produce effective autocatalytic dynamics during the cell cycle (26, 27); we have used this simple model, despite the known nonlinearity inherent to metabolic reactions. Cell division control was described by an effective restraining force. The finite duration of cell cycle, the small dynamic range of exponential accumulation ( $\sim \times 2$ ), and the imperfect nature of division cause reshuffling of the different phenotype components at division. As a consequence, rather than a pure exponential accumulation, all components accumulate with effective exponents, which vary over cycles and between components, while maintaining a positive correlation among them.

This model induces simultaneous homeostasis on all cellular components, although only one may actually affect cell division. Moreover, it results in all components exhibiting negatively correlated accumulation and initial value. This correlation is manifested as an effective restraining force with specific value that can vary depending on interactions and levels of noise. Consequently, different individual cells may have variable empirically measured values of this homeostasis parameter, including those measured for cell size.

**Predicted Consequence of Variable Homeostasis: An Independent Verification.** The variable value of the homeostasis parameter is reflected in the structure of the ACF for cell size and protein content. While averaging over the noise ensemble smooths out such oscillations, the period of individual traces can be predicted by a theory applied to the mapping model across generations (28). This prediction depends explicitly on  $\tilde{\beta}^{(k)}$  and agrees with the data, providing additional independent support for our finding of effective homeostasis parameters  $\tilde{\beta}^{(k)}$ , which vary across traces ( $k$ ) of individual cells.

**Implications to Division Regulation.** Our results taken together highlight several gaps in the current understanding of cellular homeostasis in bacteria. Why are homeostasis parameters different among cells? We have provided an illustrative model consistent with the results, where one variable controls division, and measurable correlations are indirectly induced or modulated by intracellular interactions and noise. However, there could be other sources; for example, the microenvironment in the trap could affect processes in the cell that would result in such variability. Our analysis highlights the elusive nature of the homeostasis parameter  $\beta$ , the difficulty in identifying what it represents

in terms of intracellular processes, and in particular, its relation to cell division regulation. One possibility is that division regulation is an emergent property of the cell, which arises dynamically from complex interactions. Such dynamic feedback has been suggested as an organizing principle for mesoscopic-scale systems (25). Supporting this notion is our observation that cell size is controlled to a narrow region around distinct values for each trace and the possible role of division fraction in homeostasis, which has not been investigated so far. A combination of cell size accumulation and division fraction as relevant control variables would certainly imply a global and integrated mechanism. It is also possible that some composite variables influence division more than any one of those currently measurable (34). These speculative possibilities remain to be investigated in future work.

## Methods

**Experimental Procedure and Data Processing.** Wild-type MG1655 *E. coli* bacteria were used in all experiments. Protein content was measured through the fluorescence intensity of GFP or red fluorescent protein (tdTomato) inserted into the bacteria on a high- or medium-copy number plasmid and expressed under the control of the promoter of interest. For measuring the expression level of a metabolically relevant protein, GFP was expressed from the medium copy number plasmid pZA (35) under the control of the *lac* promoter. For a metabolically irrelevant protein, GFP was expressed from the same plasmid pZA but under the control of the viral  $\lambda$ -p<sub>R</sub> promoter. For simultaneous measurement of the expression of two proteins, GFP was expressed from the high-copy number plasmid pUC19 under the control of the *lac* promoter, while tdTomato was expressed from the pZA plasmid under the control of the  $\lambda$ -p<sub>R</sub> promoter.

Cultures were grown overnight at 30 °C in LB medium (most cell size data in the text and protein expressed from  $\lambda$ -p<sub>R</sub> promoter) or in M9 minimal medium supplemented with 1 g/L casamino acids and 4 g/L lactose (M9CL) (protein expressed from the *lac* promoter and simultaneous measurements of cell size and expression from both promoters in Fig. 5, and cell size data presented in *SI Appendix*, Fig. S4). The following day, cells were diluted in the same medium and regrown to early exponential phase (OD between 0.1 and 0.2). When reaching the desired OD, cells were concentrated 10 times into fresh medium and loaded into a microfluidic trapping device (*SI Appendix*, Fig. S9-1). After trapping, fresh medium was flown continuously through it to supply nutrients.

Cells were allowed to grow in the device for dozens of generations while maintaining the temperature fixed using a made-in-house incubator. Images of the channels were acquired every 3–6 min in phase contrast and fluorescence modes using a Zeiss Axio Observer microscope with a 100 $\times$  objective. The size and fluorescence of the tracked mother cell were measured from these images using the image analysis software microbeTracker (36). These data were then used to generate traces, such as those presented in Fig. 1A (*SI Appendix*, Fig. S2-2). Growth stability of cells in the microfluidic device was verified by comparing the average division time in the first and second halves of the trace. No trend was detected in any of the experiments (*SI Appendix*, Fig. S9-2).

**Data Analysis.** Single-cell traces were analyzed using homemade MATLAB programs. Trace ACFs and linear curve fitting were calculated by their implementations in MATLAB toolboxes. Homeostasis parameters  $\beta$  (both for cell size and for protein) were estimated as the slope of the best linear fit to scatterplots such as in Fig. 1B: namely, exponential accumulation as a function of log cell size (or protein) at the start of each cycle. Using this fit,  $\xi_n$  was estimated as the difference between the data and the line. The SE of the slope in the linear fit for a single trace was estimated as  $\sigma_{\text{in}}^2(\hat{\beta}) = \sigma^2(\xi_n)/S^2$ , where  $S^2 = \sum_{n=1}^N (\ln x_n - \overline{\ln x})^2$ ,  $N$  is the number of cycles in the trace, and the trace index ( $k$ ) is omitted here for clarity. Data measured at the Jun laboratory were extracted from the webpage accompanying ref. 3 and analyzed in the same way as our data (*SI Appendix*, Fig. S10).

The temporal average over a particular trace  $k$  is represented by an overbar [e.g.,  $\overline{\ln x^{(k)}} = 1/N_k \sum_{n=1}^{N_k} \ln x_n^{(k)}$ , where  $N_k$  is the number of cycles in the trace and  $x_n^{(k)}$  are measurements in that trace]. The internal variance over the trace is computed as  $\sigma_{\text{in}}^2(\ln x^{(k)}) = 1/N_k \sum_{n=1}^{N_k} (\ln x_n^{(k)} - \overline{\ln x^{(k)}})^2$ . The corresponding SD  $\sigma_{\text{in}}(\ln x^{(k)})$  is the square root of this quantity. To characterize the entire set of traces, the average across traces is denoted by brackets:  $\langle \sigma_{\text{in}}(\ln x) \rangle = 1/M \sum_{k=1}^M \sigma_{\text{in}}(\ln x^{(k)})$ , with  $M$  being the number of traces.

The external variance quantifies the spread of temporal averages among traces and is computed as  $\sigma_{\text{ex}}^2(\overline{\ln x}) = 1/M \sum_{k=1}^M (\overline{\ln x^{(k)}} - \langle \overline{\ln x} \rangle)^2$ . We note that



the difference between averaging the logarithm and taking the log of average quantities was not significant in our data. For initial cell size, this amounted to a 2.5% discrepancy, while for division fraction, it was less than 1%. For the analysis of Fig. 4, the mapping model for individual traces (Eq. 4) was simulated with parameters mimicking the measured data: number of traces, number of cycles in each trace, etc. Internal and external variances were computed similarly for measured and simulated traces.

**Model Simulation.** We simulated the multicomponent phenotype model (Eqs. 6 and 7) for a vector of dimension  $D=50$  components. Interactions  $K_{ij}$  were independently drawn from a Gaussian distribution with mean  $1/\sqrt{D}$  and SD  $1/\sqrt{D}$ . This matrix was kept fixed over the entire ensemble of simulated traces, each with a number of cycles drawn uniformly between 30 and 250.

The “nominal” homeostasis parameter was taken to be  $\beta=0.5$ , similar to the value of the ensemble average of the experimental data. The common pivot point coordinates are taken as  $x^*=3$  and  $\phi^*=\ln 2$ . The end of each cell cycle  $n$  is determined by Eq. 2 applied to the first component  $x_{1,n}$ . At division, each component  $j$  is multiplied by an independent Gaussian variable  $f_{j,n}$  with mean 0.5 and SD  $\sigma_f=0.1$  truncated to  $(0,1)$ . The noise in division control,  $\xi_{nr}$ , is a zero mean Gaussian variable with SD  $\sigma_\xi=0.2$ .

**ACKNOWLEDGMENTS.** We are grateful to the anonymous reviewers for critical and helpful comments. We thank Omri Barak, Erez Braun, and Lukas Geyrhofer for helpful discussions. This work was supported by the US-Israel Binational Science Foundation (N.B. and H.S.) and by the National Science Foundation Grant PHY-1401576 (to M.K., H.V., J.T.N., and H.S.). H.S. was also supported by the Fulbright Fellowship.

1. Wu F, Dekker C (2016) Nanofabricated structures and microfluidic devices for bacteria: From techniques to biology. *Chem Soc Rev* 45:268–280.
2. Di Talia S, Skotheim JM, Bean JM, Siggia ED, Cross FR (2007) The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* 448:947–951.
3. Wang P, et al. (2010) Robust growth of *Escherichia coli*. *Curr Biol* 20:1099–1103.
4. Iyer-Biswas S, et al. (2014) Scaling laws governing stochastic growth and division of single bacterial cells. *Proc Natl Acad Sci USA* 111:15912–15917.
5. Yu FB, et al. (2017) Long-term microfluidic tracking of coccoid cyanobacterial cells reveals robust control of division timing. *BMC Biol* 15:11.
6. Godin M, et al. (2010) Using buoyant mass to measure the growth of single cells. *Nat Methods* 7:387–390.
7. Brenner N, et al. (2015) Universal protein distributions in a model of cell growth and division. *Phys Rev E Stat Nonlin Soft Matter Phys* 92:042713.
8. Amir A (2014) Cell size regulation in bacteria. *Phys Rev Lett* 112:1–5.
9. Schmidt-Glenewinkel H, Barkai N (2014) Loss of growth homeostasis by genetic decoupling of cell division from biomass growth: Implication for size control mechanisms. *Mol Syst Biol* 10:769.
10. Tyson JJ, Hannsgen KB (1985) Global asymptotic stability of the size distribution in probabilistic models of the cell cycle. *J Math Biol* 22:61–68.
11. Friedlander T, Brenner N (2008) Cellular properties and population asymptotics in the population balance equation. *Phys Rev Lett* 101:018104.
12. Brenner N, et al. (2015) Single-cell protein dynamics reproduce universal fluctuations in cell populations. *Eur Phys J E Soft Matter* 38:102.
13. Taheri-Araghi S, et al. (2015) Cell-size control and homeostasis in bacteria. *Curr Biol* 25:385–391.
14. Soifer I, Robert L, Amir A (2016) Single-cell analysis of growth in budding yeast and bacteria reveals a common size regulation strategy. *Curr Biol* 26:356–361.
15. Tanouchi Y, et al. (2015) A noisy linear map underlies oscillations in cell size and gene expression in bacteria. *Nature* 523:357–360.
16. Priestman M, Thomas P, Robertson BD, Shahrezaei V (2017) Mycobacteria modify their cell size control under sub-optimal carbon sources. *Front Cell Dev Biol* 5:64.
17. Banerjee S, et al. (2017) Biphasic growth dynamics control cell division in *Caulobacter crescentus*. *Nat Microbiol* 2:17116.
18. Robert L (2015) Size sensors in bacteria, cell cycle control, and size control. *Front Microbiol* 6:515.
19. Osella M, Nugent E, Cosentino Lagomarsino M (2014) Concerted control of *Escherichia coli* cell division. *Proc Natl Acad Sci USA* 111:3431–3435.
20. Grilli J, Osella M, Kennard AS, Lagomarsino MC (2017) Relevant parameters in models of cell division control. *Phys Rev E* 95:032411.
21. Kessler DA, Burov S (2017) Effective potential for cellular size control. arXiv: 1701.01725.
22. Campos M, et al. (2014) A constant size extension drives bacterial cell size homeostasis. *Cell* 159:1433–1446.
23. Shahrezaei V, Marguerat S (2015) Connecting growth with gene expression: Of noise and numbers. *Curr Opin Microbiol* 25:127–135.
24. Spiesser TW, Müller C, Schreiber G, Krantz M, Klipp E (2012) Size homeostasis can be intrinsic to growing cell populations and explained without size sensing or signalling. *FEBS J* 279:4213–4230.
25. Laughlin RB, Pines D, Schmalian J, Stojkovic BP, Wolynes P (2000) The middle way. *Proc Natl Acad Sci USA* 97:32–37.
26. Hinshelwood CN (1952) On the chemical kinetics of autolytic systems. *J Chem Soc* 1952:745–755.
27. Iyer-Biswas S, Crooks GE, Scherer NF, Dinner AR (2014) Universality in stochastic exponential growth. *Phys Rev Lett* 113:028101.
28. Newman WI (2014) Emergence of patterns in random processes. II. Stochastic structure in random events. *Phys Rev E Stat Nonlin Soft Matter Phys* 89:062113.
29. Nozoe T, Kussell E, Wakamoto Y (2017) Inferring fitness landscapes and selection on phenotypic states from single-cell genealogical data. *PLoS Genet* 13:e1006653.
30. Taniguchi Y, et al. (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* 329:533–538.
31. Larrabee KL, Phillips JO, Williams GJ, Larrabee AR (1980) The relative rates of protein synthesis and degradation in a growing culture of *Escherichia coli*. *J Biol Chem* 255:4125–4130.
32. Trötschel C, Albaum SP, Poetsch A (2013) Proteome turnover in bacteria: Current status for *Corynebacterium glutamicum* and related bacteria. *Microb Biotechnol* 6:708–719.
33. Salman H, et al. (2012) Universal protein fluctuations in populations of microorganisms. *Phys Rev Lett* 108:238105.
34. Daniels BC, Chen Y-J, Sethna JP, Gutenkunst RN, Myers CR (2008) Sloppiness, robustness, and evolvability in systems biology. *Curr Opin Biotechnol* 19:389–395.
35. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/11-12 regulatory elements. *Nucleic Acids Res* 25:1203–1210.
36. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C (2011) High-throughput, sub-pixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol Microbiol* 80:612–627.